DOCUMENT RESUME

ED 408 309 TM 026 522

AUTHOR Denton, Carolyn A.

TITLE An Evaluation of an Implementation of the Reading Recovery Program.

PUB DATE Jan 97

NOTE 26p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, January 23-25, 1997).

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Achievement Gains; Comparative Analysis; *Elementary School Students; Grade 1; *Literacy; Primary Education; *Program Evaluation; Reading Achievement; Reading Instruction; Reading Programs; Rural Schools; *Tutoring

IDENTIFIERS Iowa Tests of Basic Skills; *Reading Recovery Projects; Texas

ABSTRACT

The effectiveness of the Reading Recovery Program, a literacy intervention for at-risk first-grade students, as it was implemented in 1995-96 in a small rural school district in Texas was studied. Reading Recovery consists of daily individual literacy tutoring sessions taught by a specially trained teacher, and emphasizes active engagement of the child, instruction in use of literacy strategies, instructional scaffolding, and teaching for independence. Quantitative data was collected over a 9-month period, comparing the performance and rate of improvement of treatment groups of four and six students with those of a comparison group of nine students who received instruction in small literacy development groups. The number of students in the comparison group decreased as students moved into the second treatment group. The achievement of treatment and comparison students was compared to that of all 33 first graders in the school. Student performance was measured using an informal survey of oral reading fluency, accuracy, and comprehension, and the reading components of the Iowa Test of Basic Skills. The program seemed to meet some of its goals, but fell short of others. There is evidence that the program helped students to achieve an accelerated rate of growth in their literacy learning and to read accurately and with average comprehension. The treatment group students did not achieve expected normative benchmarks in oral reading fluency or the norm-referenced measure, and they did not seem to attain the average level of performance of their peers on literacy measures. (Contains 33 references.) (SLD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED 408 309

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
CAROLYN A. DENTON
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

# An Evaluation of An Implementation of the Reading Recovery Program

Carolyn A. Denton
Texas A&M University

BEST COPY AVAILABLE

Paper presented at the 20th annual meeting of the Southwest Educational Research Association, January 23-25, 1997, Austin, Texas

TM026522

An Evaluation of An Implementation of the Reading Recovery Program

Abstract

This study assessed the effectiveness of the Reading Recovery program, a literacy intervention for at-risk first grade students, as it was implemented in 1995-96 in a small rural school district in Texas. Quantitative data collected over a nine-month period were analyzed. The performance and rate of improvement of two groups of students in the Reading Recovery program, Treatment Group 1( n=4) and Treatment Group 2 (n=6), were compared to that of students on a waiting list for the program, who formed a Comparison Group (n=9). These students received instruction in small literacy development groups. The number of students in the Comparison Group decreased as students on the waiting list became part of the second Treatment Group. The achievement of Treatment and Comparison students was compared to that of a Peer Group consisting of all first grade students in the school (n=33), and to normative benchmarks. The progress of individual students was examined in more detail as case studies.

The evaluation examined the extent to which the program met its stated goals. It described the success rate of the program, the duration of treatment for program participants, and the text reading levels of Treatment Group students at the end of grade one. The study compared the rate of improvement of individual students in the Comparison Group before and after they entered the Reading Recovery program. To answer the research questions of this study, student performance was measured using components of the Observation Survey of Marie Clay (Clay, 1993a), an informal inventory of oral reading fluency, accuracy, and comprehension, and the reading components of the Iowa Test of Basic Skills.

The evaluation found that the program seemed to meet some of its goals, but fell short of meeting others. There is evidence that the program helped students to achieve an accelerated rate of growth in their literacy learning and to read accurately and with average comprehension. However, the Treatment Group students did not achieve expected normative benchmarks in oral reading fluency or on the norm-referenced measure, and they did not seem to attain the average level of performance of their peers in literacy measures.

An Evaluation of An Implementation of the Reading Recovery Program

This study assessed the effectiveness of the Reading Recovery (RR) program, a literacy intervention for at-risk first grade students, as it was implemented in 1995-96 in a small rural school district in Texas. The RR program consists of daily individual literacy tutoring sessions taught by a specially trained teacher, emphasizing active engagement of the child in reading and writing activities, instruction in the use of strategies on text (including self-monitoring and self-correction), observation of the child's literacy behaviors to guide instruction, instructional scaffolding, and teaching for independence.

## Problem and Need

The evaluation examined the extent to which the program met its stated goals. The program goals are (a) to accelerate the rate of progress of individual first grade students who currently perform at the lowest levels in their respective classes in literacy tasks, (b) that at least 75% of the program participants will reach performance criteria established by the program for discontinuation of tutoring, (c) that these levels will be achieved after an average of 60 lessons, not including the 10 sessions of "roaming in the known" (Clay, 1993b), (d) that these students will read "on grade level" by the end of grade one (in RR, a text level 16), and (e) that these children will perform literacy tasks within the average range of performance of all first grade students in the school at the end of grade one. There was a need for both formative and summative data.

### General Approach to Meeting Need

The evaluation examined the success of individual students in meeting the program goals, and it studied the number of lessons required to meet success criteria. The performance and rate of improvement of two groups of students in the RR program, Treatment Group 1 (TG1) (n=4) and Treatment Group 2 (TG2) (n=6), were compared to that of students on a waiting list for the program, who formed a Comparison Group (n=9); these students received instruction in small literacy development groups taught by the RR teacher. Some of the students in these literacy groups entered the Reading Recovery program at different times during the year, as

openings became available, forming TG2. Their rates of improvement before and after they entered the program were compared. In addition, the achievement levels of the TG students were compared to the average achievement level of the Peer Group consisting of all first grade students in the school (including TG and CG students) (n=33). The purpose of this comparison was to determine whether students in RR performed in the average range for students in this school at the end of grade one. As a further evaluation of the program's effectiveness, the progress of individual students was examined in greater detail, as case studies. The research was conducted by the district's RR teacher.

Research Questions

The following questions guided the study:

1. What percentage of TG students were successfully discontinued from the program? Which individual students from the TG were successful? Were treatment children able to achieve success after approximately 60 RR lessons, not including the period of "roaming in the known" (Clay, 1993b)? Did these students achieve a text reading level of at least 16 (grade 1.9) at the end of grade one?

2. Did students in TG1 show a greater rate of improvement over time (during treatment) on measures of oral reading accuracy and writing vocabulary than the students in the Comparison Group?

3. Did students who began the study as members of the Comparison Group and later entered the TG demonstrate a greater rate of improvement on the above measures after entering the program?

4. Did students in TG1 achieve a higher level of performance on a measure of oral reading fluency, accuracy, and comprehension administered in January than did students in the Comparison Group?

5. Did the students in TG1 and 2 achieve an accuracy level of at least 90% (considered instructional level in RR) and a comprehension level of at least 70% on a similar measure administered in May? Did TG1 and 2 students attain levels of oral reading fluency considered

normative for the beginning of grade two (Hasbrouck & Tindal, 1992) on the measure administered in May? Did students who were considered successful in the program attain the above goals?

6. Did students in TG1 and 2 attain scores in the average range of performance of the Peer Group on tests of oral text reading, writing vocabulary, and accurate recording of phonemes in a dictated passage administered at the end of grade one? Did students who were successfully discontinued from the program achieve these levels?

7. Did students in TG1 and 2 attain scores in the average range of the Peer Group on a norm-referenced test of reading comprehension administered near the end of grade one? Did students who were successfully discontinued from the program achieve these levels?

8. What were the particular circumstances affecting the progress of certain individuals, successful and unsuccessful, in TG1 and 2, and in the Comparison Group?

## Method

### Participants

The subjects of this study were identified by their classroom teachers as performing in the lowest 50% of their respective classes in literacy tasks. All students so identified were subsequently tested by the RR teacher, using Clay's Observation Survey (Clay, 1993a). The lowest-performing students on this measure were selected by a RR teacher leader. There was one exception to this rule of accepting the lowest performing students for the treatment. One child with limited English proficiency was judged by the classroom teacher, the English as a Second Language (ESL) teacher, and the RR teacher to have insufficient facility with the English language to benefit from the program at the beginning of the year. This child was served later in the year, and was part of TG2. It is also the policy of RR not to serve children who are repeating grade one. Two retainees who would have qualified for treatment remained in the literacy group at this time; one was provided the RR program unofficially as part of TG2. Once selected for the program, students remained in treatment until tested and discontinued (successfully) or dismissed (not successfully) from the program by the teacher leader. When a

child was removed from the treatment, other children on the waiting list were again evaluated using the Observation Survey, and the child who currently had the lowest performance on this measure was chosen to fill the program vacancy. The Peer Group consisted of all first grade students in the school at the end of the school year, including TG1, TG2, and Comparison Group students. This group was composed of 33 students, 17 males and 16 females. The following table illustrates pertinent information about the TG and Comparison Group students.

Table 1

Characteristics of TG and Comparison Group Students

| Student | Sex | Age at entry | Ethnicity | Entry | Exit | Other |
|---|---|---|---|---|---|---|
| | | | TG1 | | | |
| TG1-A | M | 7 | AA | 8/29/95 | 2/27/96 | SPED |
| TG1-B | M | 6 | W | 8/29/95 | 4/3/96 | |
| TG1-C | M | 6 | Hisp. | 8/29/95 | 3/8/96 | LEP |
| TG1-D | F | 6 | W | 8/30/96 | 3/4/96 | |
| | | | TG2 | | | |
| TG2-A | F | 6 | W | 3/5/96 | 5/24/96 | |
| TG2-B | F | 7 | Hisp. | 3/5/96 | 5/24/96 | |
| TG2-C | M | 7 | W | 4/4/96 | 5/24/96 | |
| TG2-D | M | 7 | Hisp. | 3/18/96 | 4/22/96 | LEP |
| TG2-E | M | 8 | Hisp. | 3/6/96 | 5/20/96 | LEP, R |
| TG2-F | F | 7 | W | 4/25/96 | 5/24/96 | |
| | | | Comparison Group, not part of TG2 | | | |
| CG-A | F | 6 | W | 8/23/95 | 1/25/96 | |
| CG-B | F | 6 | W | 8/23/95 | 2/15/96 | |
| CG-C | M | 7 | Hisp. | 9/19/95 | 5/2/96 | LEP, R |

AA: African American; W: White, not Hispanic; Hisp.: Hispanic
LEP: Limited English Proficient; SPED: Served by Special Education; R: Retained in grade 1

Context

The RR program was in its first year of implementation at the school at the time of this study. There was one RR teacher in the school. The school, which served Kindergarten through sixth grade, was relatively small, with a total of 212 pupils. The population of the school in 1994 was 69.2% white, 27.1% Hispanic, 3.3% African American, and .5% "other". The economically disadvantaged students in the same year were 38.3% of the total campus population, and 6.1% of the students had limited English proficiency. This school was located in a small rural community with a population of approximately 1,100. Much of the population was employed in agriculture-related activities. At the time of this study, there were two first grade classes in the school; Classroom One had 16 students and Classroom Two had 17.

Instruments and Procedures

The following instruments were used to collect data to answer the research questions:

1. Program records were examined to determine the length of treatment and the individual achievement of students within the program.

2. Achievement or rate of improvement over time in oral reading accuracy, accurate recording of phonemes in a dictated passage, and writing vocabulary were assessed using Tests 4 (Writing Vocabulary), 5 (Dictation), and 6 (Text Reading Level) of the Observation Survey of Marie Clay (Clay, 1993a). These were administered by the school's RR teacher, another trained RR teacher, or a RR teacher leader.

In Test 4 (Writing Vocabulary), students are asked to write all the words they can in a ten minute period. The test administrator prompts the student by suggesting categories of words, such as color words, number words, names of animals, etc. This prompting is not standardized, but suggestions for prompts are provided to the administrator. The score is the total number of words written which are spelled correctly (Clay, 1993a).

In Test 5 (Dictation), students hear two sentences read by the test administrator. These can be repeated slowly, word by word. The student is asked to write what he/she hears. The

score is the total number of phonemes recorded correctly (correct spelling is not required). There is a total of 37 phonemes in each passage (Clay, 1993a).

In Test 6 (Text Reading Level), students read actual text passages which have been leveled according to their difficulty. Students are presented with passages of increasing difficulty until they are unable to read two successive passages with at least 90% accuracy. The score is the highest level read with at least 90% accuracy.

The Observation Survey was chosen as one instrument for measurement of student growth because it is proximal to the program, and because it is the instrument by which RR students and programs are universally evaluated. However, this instrument has certain shortcomings in reliability. The reliability of Test 4, Writing Vocabulary, is highly questionable. In this test, the test administrator is allowed a great deal of discretion in deciding when and how to prompt for words which may be known by the child. The frequency and nature of prompting could create high variability in administration. Test 5, Dictation, has a ceiling effect which limits its effectiveness as a measure of student growth, particularly at the end of grade one. Test 6, Text Reading, has been questioned because it is based on a scale which is not divided into strictly equal intervals of difficulty (Center et al., 1992).

3. Proficiency in oral reading fluency, accuracy, and comprehension was assessed using an informal reading inventory (IRI). It consisted of three stories with comprehension questions administered in January and three different but equivalent stories administered in May. The reading selections were pilot tested with three first grade students who were not part of the current study, in order to ensure appropriateness and equivalence. Reading material was taken from widely accepted grade one published materials, with introductions and comprehension questions developed by the researcher. The assessment was conducted by the RR teacher.

4. The reading portion of the Iowa Test of Basic Skills (ITBS), Form L, Level 7, Complete Battery (Hoover et al., 1993) was administered by the classroom teachers in April to provide norm-referenced scores in reading comprehension and vocabulary.

The Observation Survey components were administered as students entered and left the program, and were administered to all first grade students at the end of the school year. The informal reading inventory was administered to Treatment and Comparison Group students in January and in May. The Iowa Test of Basic Skills was administered to all first grade students in April.

## Results and Discussion

This study evaluated the effectiveness of the Reading Recovery (RR) program, as it was implemented during the 1995-96 school year in a small rural school district in Texas. The purpose of the study was to determine whether the program met its stated goals. This discussion will examine the extent to which the program met each of its goals.

### Accelerated Rate of Growth

Goal 1: To accelerate the rate of progress of individual first grade students who currently perform at the lowest levels in their respective classes in literacy tasks.

The central claim of the Reading Recovery program is that it helps students accelerate the rate of their literacy learning so they can "catch up" with the average students in their respective classes. In comparing the rate of growth of TG1 to that of the Comparison Group, and in comparing the rate of improvement of TG2 before and after they entered the program, there is an indication that the RR treatment was effective in meeting this goal. The following tables illustrate the comparison of rates of improvement over time between the treatment and comparison groups, and between the baseline and intervention phases for TG2.

Table 2

Improvement in Oral Reading Accuracy (levels per week) and Writing Vocabulary (words per week): TG1 and Comparison Group

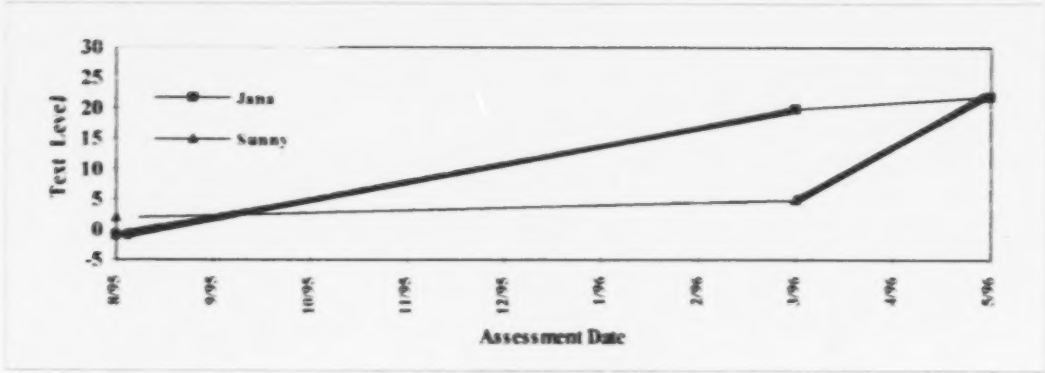
| | TG1 (n=4) | | Comparison Group (n=9) | |
|---|---|---|---|---|
| | M | SD | M | SD |
| Levels per week | .52 | .29 | .26 | .19 |
| Words per week | 1.76 | .36 | 1.38 | .46 |

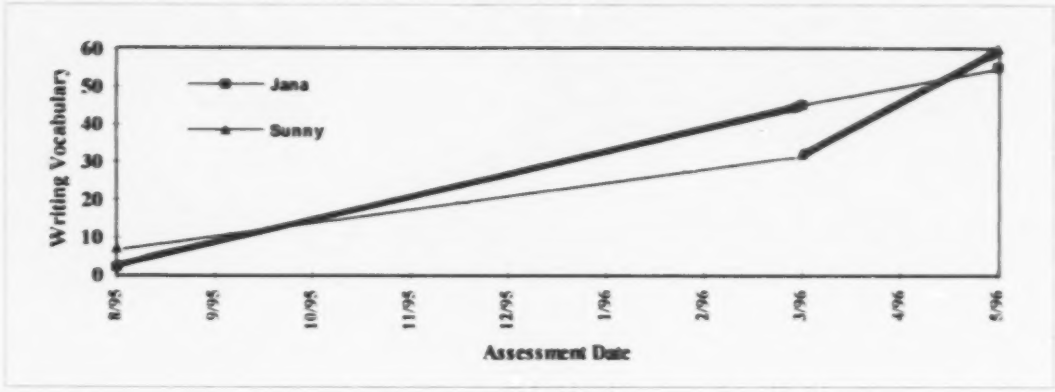
Table 3

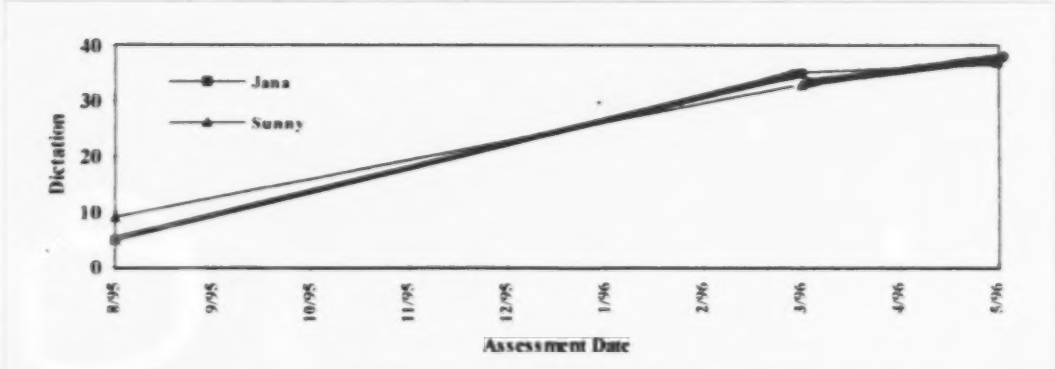
Improvement in Oral Reading Accuracy (levels per week) and Writing Vocabulary (words per week) Before and After Treatment: TG2

| | Before Entry | | After Entry | |
|---|---|---|---|---|
| | M | SD | M | SD |
| Levels per week | .18 | .04 | 1.05 | .66 |
| Words per week | 1.22 | .47 | 2.03 | .72 |

Case study data seemed to support the achievement of accelerated growth by some treatment students. A comparison was made between matched pairs of students from the same classrooms and with similar backgrounds and pretest performance levels. An examination of the Observation Survey scores of three matched pairs of children showed that, in every case, the rate of progress in writing vocabulary and text reading level was markedly increased while students were in treatment. The following graphs compare growth rates on OS measures between two matched pairs of students, while in and out of treatment. Figure 1 compares rates of growth of two white females, Jana (TG1-D), a member of TG1, and Sunny (TG2-A), a member of TG2. Figure 2 compares rates of improvement rates of two Hispanic males, both with LEP, and both retained in and repeating grade one. One (Noel: TG2-E) was served unofficially in RR and received an incomplete program at the end of the school year, and one (Juan: CG: C) was not served.

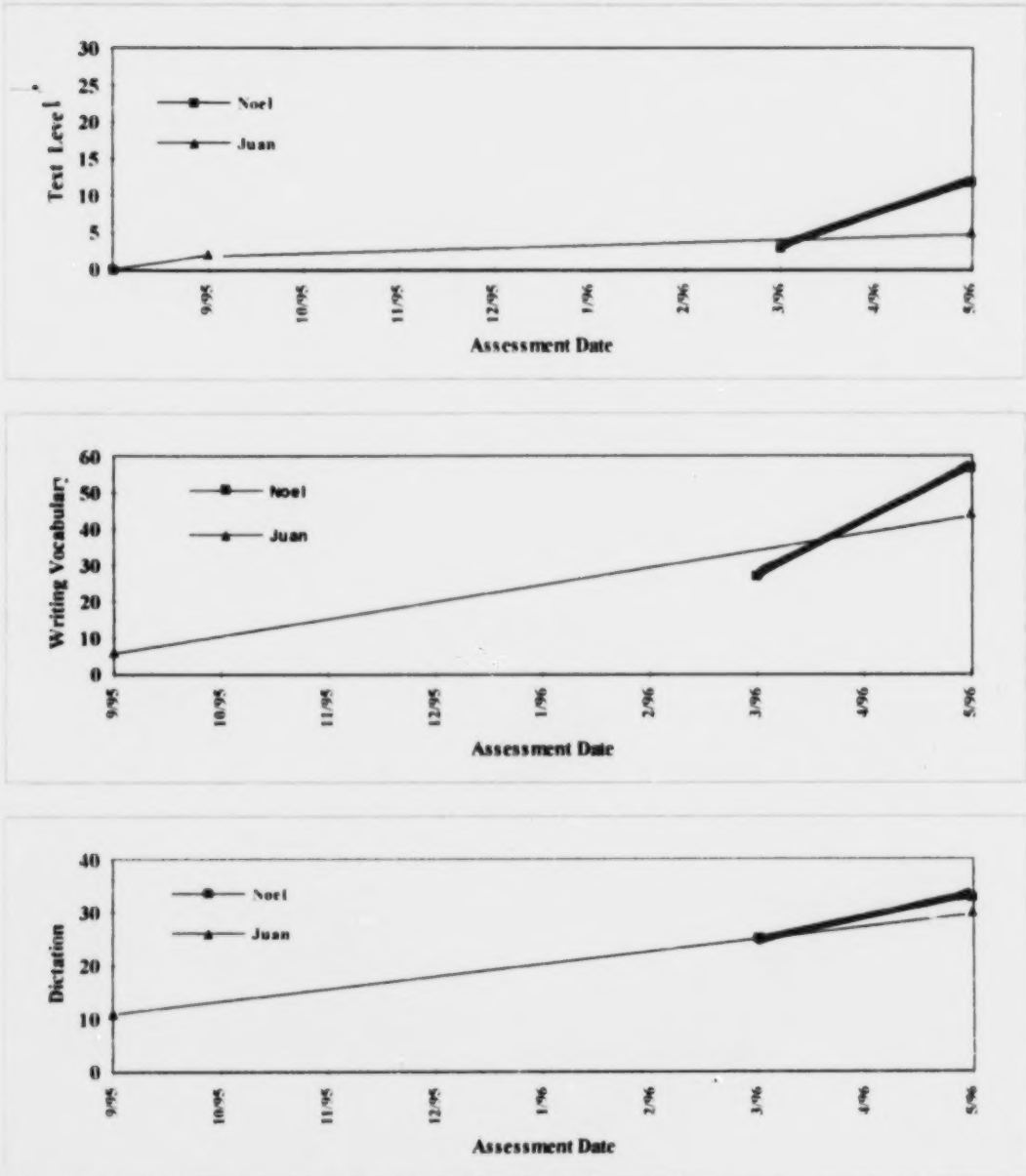



Note. Bold lines denote progress during treatment periods.

Figure 1. Assessments of text level and writing vocabulary showed that each student had a higher rate of growth during the period in which they were in treatment. The two girls had similar rates of growth in the dictation measure, regardless of treatment status.




Note Bold lines denote progress during treatment periods.

Figure 2. Assessments indicate that the two boys had similar performance levels on these tasks after one year and six months in a first grade program. Noel demonstrated an increased rate of growth in all areas after the initiation of tutoring.

Children in TG2 seemed to make very little progress with their regular classroom instruction before the individual treatment began, but their progress was accelerated after entry into tutoring. This strongly supports the use of RR with these students. Students who entered the program in March or April may have been frustrated by their classroom programs. They were unable to perform successfully the tasks that were required of them, and they seemed to suffer from low self-concept in regards to their reading and writing ability. One of the central philosophies on which RR is based is that children need to be assessed to discover what they can do independently, then instructed on a level on which they can succeed with support from the teacher, and that their individualized programs should be built on their strengths, not their deficits (Clay, 1991; 1993b). These components of the program may have supported accelerated growth in TG1 and 2 students. This pattern of accelerated growth was noted both in students who were native English speakers, and in students who had limited English proficiency, and it was observed in a student who had been retained in first grade.

However, acceleration of progress was not noted in one student who was served by Special Education (having Attention Deficit Hyperactivity Disorder), or in one student with limited English proficiency. In both of these cases, students did not seem to adopt the reading strategies promoted by the program. The first student seemed to rely on visual memory for text, and he did not appear to use phonetic analysis or context cues to read unknown words. He did not seem to be monitoring his reading for meaning. It was also noted that he had been diagnosed with vision problems. It is difficult to determine how much this might have interfered with his processing. The second student had limited experience with the English language. He was served in RR for only a short time, and he might have made more progress had he been offered treatment for a more extended period. He did not seem to adequately use the strategies which RR teaches.

RR fosters accelerated growth only when children adopt and employ the reading strategies which are central to the program. There may be a variety of reasons why some children do not use these strategies, including conditions within the child, conditions within the

teacher and her manner of teaching, and/or conditions inherent in the relationship between the child and the teacher. RR puts heavy emphasis on the responsibility of the teacher to observe and accommodate the needs of the child, and on the power of the social relationship between child and teacher (Clay, 1993b). Some implications are that students need to be provided adequate time in treatment, and that the RR teacher should carefully examine her teaching decisions and her relationship with each child for factors which could impede student growth. The teacher should closely analyze her written records, and she might also make audio recordings of lessons in order to identify factors which may fail to stimulate student progress.

Success Rate of the Program

Goal 2: That at least 75% of the program participants will reach performance criteria for discontinuation of tutoring established by the program.

A total of ten individual first grade children were served by the RR program during the target period. Of these, seven (70%) were successfully discontinued, or considered successful in the program. Of the three who were not successfully discontinued, one received a full program (at least 60 formal lessons), one was removed because of lack of progress, and one was served unofficially and had received an incomplete program when the school year ended. Of the children who were served officially as RR children (n=9), 78% were successful in the program. This rate of success in the program meets the goal established in this school, and is consistent with discontinuation rates published in several studies (Clay, 1993b; Pinnell, 1989; RR at OSU, 1994; TWU RR Staff, 1995). These studies have shown that RR is not a successful intervention for every child who is having problems learning to read in first grade. The percentage of students for whom the program was effective in the current evaluation is about what would have been expected, based on previously published literature.

Duration of Treatment

Goal 3: That these levels will be achieved after an average of 60 lessons, not including the 10 sessions of "roaming in the known" (Clay, 1993b).

Students in TG1 received an average of 98 lessons. This was clearly longer than the goal of discontinuation after approximately 60 lessons. Students in TG2, however, received an average of 24 lessons. The overall mean number of lessons for both groups was 54, below the goal of 60. Students in TG1 gained an average of .52 text levels per week while in treatment, while the children in TG2 gained an average of 1.05 levels per week.

Students who began their first grade year in RR had minimal literacy skills, a general lack of awareness of print, and little previous experience with text. These students required an extended intervention in order to reach average performance levels. One factor which may have affected the duration of treatment for this group is the fact that the RR teacher was in training during the target year; she would be expected to become more effective as the year-long training program progressed.

Students in TG2 entered the program with a base of basic literacy skills on which to build. The writing vocabulary and dictation levels of most TG2 students were much higher at the time of entry into RR than those scores had been for TG1 students when they started tutoring. It could be inferred that the TG2 students had a basic knowledge of sound-letter correspondence and of some spelling conventions. However, the text reading levels of the TG2 students at the time of program entry were below expected levels for their grade, and their classroom teachers reported that they were not successful in their classroom reading programs. Students in this group made the most rapid progress within the RR program. Another factor affecting the number of lessons required for TG2 students was their prior participation in literacy groups. These children were already familiar with the RR teacher and some procedures which were common to the two settings.

Reading Level of Tutored Students

Goal 4: That these students will read "on grade level" by the end of grade one (in RR, a text level 16).

Of the ten TG students, six achieved the program goal of accurate reading of at least level 16 material at the end of grade one, considered equivalent to a 1.9 basal level. Two of the

students who did not achieve this goal had been removed from treatment because it was not successful for them. One had not received a complete program at the end of the school year, and may have made more progress given the time and opportunity. One student who had been successfully discontinued from TG1 did not achieve grade level reading at the end of grade one. The program seems to have failed in its claim to provide this student with a "self-extending system" which allows children to continue to improve independently after discontinuation (Clay, 1993b). It could be assumed that treatment may have been discontinued for this child too soon, before he was secure in his reading skills. It should be noted that this child had limited English proficiency and was from a classroom environment in which there was little opportunity provided for independent reading or writing. Implications for the RR program are that better coordination with students' classroom programs might increase the effectiveness of the intervention, and that students with limited English proficiency may need additional support after tutoring has been discontinued.

The ability of students to read at expected levels for their grade was observed with multiple measures. The components of the Observation Survey which produced the above results were proximal to the program itself. The purpose of the administration of an IRI was to measure student achievement with a more distal tool. In program evaluations of RR, program students typically perform very well when measured with Clay's survey, but more distal measures show weaker program effects (Donley & Baenen, 1993; Ramaswami, 1994; Salinas et al., 1993). This seems to be the case in the current research as well, and may be the result of the failure of RR to adequately emphasize the transfer of skills and behaviors from the tutoring environment to other situations. On a measure of oral reading fluency administered in January, Comparison Group students were able to read more fluently and answer a higher percentage of comprehension questions correctly than were TG students. The TG read with slightly more accuracy, however.

Table 4

Results of January IRI: TG1 and Comparison Group (CG)

| | WCPM* | Accuracy | Questions Correct |
|---|---|---|---|
| TG1 (n=4) | | | |
| M | 7 | 57% | 35% |
| SD | 1.7 | 13 | 22 |
| CG (n=8**) | | | |
| M | 7.4 | 52% | 40% |
| SD | 4.4 | 17 | 21 |

*WCPM: Words read correctly per minute

**One member of CG was not assessed. The comprehension score was not counted for one addition member of CG because the child was told numerous words during the reading. Many of these words told were key to the meaning of the passage. Her score was judged invalid.

The lack of fluency of the RR students may be the result of the type of reading strategies the program supports. RR students are taught to self-monitor and to repeat words or phrases of text if they do not sound right, look right, and make sense, with a goal of self-correcting errors (Clay, 1993b). This would promote accurate reading, the primary goal of this emergent literacy program, but it would not foster fluent reading of unfamiliar text. The RR program does consider fluency a goal for students, but it defines fluency in terms of phrased and expressive reading, rather than overall reading rate (Clay, 1993b). As RR students become more competent readers, however, their general fluency should increase, as fewer errors would require fewer repetitions. After four months in treatment, the TG did not perform in oral reading fluency as well as students who had had no intervention.

On a similar measure administered in May, the mean performance levels of the TG still fell short of normative benchmark goals.

Table 5

Results of May Informal Reading Inventory: TG1, TG2 and Successfully Discontinued Students

| | WCPM | Accuracy | Questions Correct |
|---|---|---|---|
| TG1 & 2 (n=10) | | | |
| M | 21 | 84% | 66% |
| SD | 7.5 | 13 | 37 |
| SDS* (n=7) | | | |
| M | 21 | 90% | 88% |
| SD | 7.6 | 4 | 12 |
| Norms or Benchmarks | 53 | 90% | 70% |

*SDS: Successfully discontinued students

It is interesting that the subgroup of successfully discontinued students performed no better in oral reading fluency than did the full TG cohort. Both groups fell short of the norm of 53 WCPM established for the 50th percentile at the beginning of grade two, and were slightly lower than the norm of 23 WCPM established for the 25th percentile at this level (Hasbrouck & Tindal, 1992). The TG did not demonstrate the ability to read as fluently as would be expected for students of their age and grade.

When this group's performance on the IRI was analyzed for reading accuracy, however, students fared much better. The accuracy rate of the entire cohort was 84%, just under the goal of 90%, an indication of instructional reading level (Clay, 1993a). When the scores of students who had been successful in the RR program were considered, this subgroup read the text at an average accuracy rate of 90%. This would indicate that students whom the program staff judged to be successful readers and writers could read grade one non-program material at at least an instructional level. It is also noteworthy that successfully discontinued students answered an average of 88% of the comprehension questions correctly in this test administration. They

seemed to read accurately and to monitor what they read for meaning. While showing program weakness in promoting oral reading fluency, this measurement does indicate success in promoting accurate reading and text comprehension. These results may indicate the need for program personnel to follow the RR treatment with instruction designed to promote oral reading fluency.

Comparisons to Peer Group

Goal 5: That these children will perform literacy tasks within the average range of performance of all first grade students in the school at the end of grade one.

A primary stated goal of RR is to bring the performance levels of the lowest-achieving children into the range of average performance for their class. Target students in this study did not appear to meet that goal. The mean performance of program children was below the average range of their peers in Text Reading Level, Writing Vocabulary, and Dictation measures. When students who were successful in the intervention were considered as a subgroup, their mean performance was only slightly below the average range for Text Reading Level, and was within that range in Dictation. Both of these tests had ceiling effects, however, and the actual range of ability of the Peer Group may have been higher, but unmeasured by the instrument used.

It cannot be conclusively stated that RR children performed as well as the average of the Peer Group. However, an evaluation of oral reading accuracy of a random sample of 273 non-treatment students from Texas RR schools conducted in the spring of 1995 showed that non-program Texas students read at an average text level of 18.97, producing an average range (using .5 SD above and below the mean) of 14.54 to 23.40 (TWU RR Staff, 1995). The TG mean of 15 falls just within this range, and the successfully discontinued subgroup mean of 18.29 is almost equal to the state-wide mean of 18.97. The TG mean scores also fall within the average range of the Texas non-program means for both Writing Vocabulary and Dictation. The program could be said to produce students who read and write accurately in the average range for the state of Texas, but not in the average range for their own school.

Table 6

Observation Survey Measures: TG, Peer Group, and Successfully Discontinued Students

| | Text Reading Level | Writing Vocabulary | Dictation |
|---|---|---|---|
| Peer Group (n=33) | | | |
| M | At least 22.88* | 71.3 | 35.8 |
| SD | 8.1 | 18.7 | 1.9 |
| Average performance ** | 18.82 - 26.95 | 61.9 - 80.6 | 34.8 - 36.7 |
| TG1 & TG2 (n=10) | | | |
| M | 15 | 53.4 | 34.6 |
| SD | 6.18 | 7.4 | 2.1 |
| SDS (n=7) | | | |
| M | 18.29 | 56 | 35.6 |
| SD | 3.15 | 5.5 | 1.6 |

* Testing of Peer Group students was arbitrarily stopped on level 30. This mean may have been higher if testing had continued on higher levels. ** In RR literature, performance of tutored students is compared to an "average range" of performance of a peer group, created by adding and subtracting the amount of .5 SD to the mean peer group score.

When the performance of the students was measured using a norm-referenced standardized test, program students clearly fell short of the goal of performance within the average range of their own school peers. The percentile rankings of TG and successfully discontinued students were well below the average percentile ranking of their peer group. It should be noted, however, that the ITBS was administered in April; the TG2 students had not received the full benefit of their RR program at that time. Nevertheless, this analysis of test data indicates that the program was not successful in raising student performance to average levels on this measure, an instrument unrelated to the program.

Table 7

TG Performance and Peer Group Performance in National Percentile Ranks: Iowa Test of Basic Skills (Reading)

| | Vocabulary | Comprehension | Total Reading |
|---|---|---|---|
| PG (n=34) | | | |
| M | 55 | 58 | 55 |
| SD | 31 | 25 | 30 |
| TG1 & 2 (n=10) | | | |
| M | 24 | 34 | 22 |
| SD | 24 | 11 | 19 |
| SDS (n=7) | | | |
| M | 30 | 38 | 27 |
| SD | 26 | 8 | 20 |

The relatively low student performance on standardized test materials has important implications for program implementation. Linda Dorn (1996) suggests that some RR students may be stimulated to perform by environmental factors such as the presence of the RR teacher, the setting of the RR room, or the format of the reading material presented to them. Children may not realize that they can use what they have learned in many different situations, or they may not have internalized the behaviors well enough to perform in the absence of environmental stimulation. Fostering independence in the child is strongly emphasized by Clay. She maintains that the role of the RR teacher is to aid in the construction and development of inner control by the child himself (Clay, 1991, 1993b). Dorn (1996) suggests that RR teachers should select materials and structure activities with the clear goal of promoting flexibility and transfer of skills and behaviors, and then observe the child keenly to find out whether transfer is taking place. This concept has important implications for the implementation of RR in the current study, and

in many other situations. There is also an indication that RR tutoring should be followed by instruction emphasizing silent reading comprehension.

## Conclusions

The study concluded that the Reading Recovery program, as it was implemented in this district, did appear to have met some of its goals, but it seemed to have fallen short of meeting others. The evaluation supports the continuation of the program, but it does have implications for improved implementation within this district, and for the RR program in general. The most important of these are increased coordination between the classroom and RR programs of tutored students, increased support for program children with limited English proficiency, and greater emphasis on teaching for transfer of learning from the tutoring situation to other contexts. There is also evidence to support the following of RR tutoring with instruction which emphasizes oral reading fluency and silent reading comprehension.

This study was limited by the small sample size, by the effect of regression to the mean, by potential unintentional observer bias, by the effect of dedicated time in the program, and by certain shortcomings of the observation instruments. These limitations were unavoidable in the setting in which the research was carried out. As more and more program students are monitored, a larger body of data will accumulate. If possible, future evaluation should be conducted by persons who do not have a stake in the program, and the use of multiple measures should be continued. Studies should follow program student progress longitudinally to monitor the maintenance of program effects as the students progress through school. In addition, future study should focus on changes in program implementation, such as increased emphasis on coordination with the classroom programs and transfer of behaviors to different environments, in order to determine whether these changes increase program success.

References

Allington, R.L. & McGill-Franzen, A. (1990). Children with reading problems: How we wrongfully classify them and fail to teach many to read. ERS Spectrum, 8, 3-9.

Center, Y., Wheldall, K., & Freeman, L. (1992). Evaluating the effectiveness of Reading Recovery: a critique. Educational Psychology, 12, 263-274.

Clay, M.M. (1985). The early detection of reading difficulties. Auckland, New Zealand: Heinemann.

Clay, M.M. (1991). Becoming literate: The construction of inner control. Portsmouth, NH: Heinemann.

Clay, M. M. (1993a). An observation survey of early literacy achievement. Portsmouth, NH: Heinemann.

Clay, M.M. (1993b). Reading Recovery: A guidebook for teachers in training. Portsmouth, NH: Heinemann.

Donley, J. & Baenen, N. (1993). A study of the long-term effectiveness of the Reading Recovery program. (E&R Report No. 93.09A). Wake County, NC: Wake County Public School System. (ERIC Document Reproduction Service No. ED 359 197).

Dorn, L. (1996, October). Teaching for transfer: Promoting independent activity across changing circumstances. Paper presented at the annual Reading Recovery Conference presented by Texas Woman's University, Dallas-Fort Worth Airport, TX.

Dyer, P.C. (1992). Reading Recovery: A cost-effectiveness and educational-outcomes analysis. ERS Spectrum, 10, 10-19.

Edmark Reading Program, 2nd Ed. (1992). Redmond, WA: Edmark Corp.

Escamilla, K. & Andrade, A. (1992). Descubriendo La Lectura: An application of Reading Recovery in Spanish. Education & Urban Society, 24, 212-226.

Hasbrouck, J. & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. Teaching Exceptional Children, 24, 41-44.

Hiebert. E.H. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? Educational Researcher, 23, 15-25.

Hoover, H.D., Hieronymus, A. N., Frisbie, D.A., & Dunbar, S.B., (1993). Iowa Tests of Basic Skills, Complete Battery, Form K, Level 7. Chicago: Riverside Publishing Company.

Ihnot, C. (1992). Read Naturally. St. Paul, MN: Read Naturally.

Iverson, J.A. & Tunmer, W.E. (1993). Phonological processing skills and the Reading Recovery program. Journal of Educational Psychology, 85, 112-125.

Lyons, C.A. (1989). Reading Recovery: An effective early intervention program that can prevent mislabeling children as learning disabled. ERS Spectrum, 7, 3-9

Lyons, C.A. (1991). Reading Recovery: A viable prevention of learning disability. Reading Horizons, 31, 384-408.

Lyons, C.A. & Beaver, J. (1995). Reducing retention and learning disability placement through Reading Recovery: An educationally sound, cost-effective choice. In R.L. Allington & S. A. Walmsley (Eds.), No quick fix: Rethinking literacy programs in America's elementary schools (pp. 116-136). New York: Teachers College Press.

North American Reading Recovery Council (NARRC). (1993). Guidelines and standards for the North American Reading Recovery Council (2nd ed.). Columbus, OH: The Ohio State University.

Organisation for Economic Cooperation and Development (OECD) & Statistics Canada (1995). Literacy, economy, and society: Results of the first international adult literacy survey (Report No. ISBN-92-64-14655-5). Paris, France & Ottawa, Ontario: Author. (ERIC Document Reproduction Service No. ED 388 881).

Pinnell, G.S. (1989). Reading Recovery: Helping at-risk children learn to read. Elementary School Journal, 90, 161-183.

Pinnell, G.S., Lyons, C.A., DeFord, D.E., Bryk, A.S. & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. Reading Research Quarterly, 29, 9-39.

Ramaswami, S. (1994). The differential impact of Reading Recovery on the achievement of first graders in the Newark School District, 1991-1993. Newark, NJ: Newark Board of Education, NJ Office of Planning, Evaluation and Testing. (ERIC Document Reproduction Service No. ED 374 180).

Rasinski, T. (1995). Commentary on the effects of Reading Recovery: A response to Pinnell, Lyons, DeFord, Bryk, and Seltzer. Reading Research Quarterly, 30, 264-270.

Reading Recovery Program at OSU. (1994). National Diffusion Network executive summary: Reading Recovery 1984-1993. Columbus, OH: OSU.

Ross, S. M., Smith, L. J., Casey, J, & Slavin, R.E. (1995). Increasing the academic success of disadvantaged children: An examination of alternative early intervention programs. American Educational Research Journal, 32, 773-800.

Salinas, A., Williams, I., & Kohler, L. (1993). Reading Recovery program evaluation report 1992-93. McAllen, TX: McAllen Independent School District. (ERIC Document Reproduction Service No. ED 366 938).

Shanahan, T. & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. Reading Research Quarterly, 30, 958-996.

Shepard, L.A. & Smith, M.L. (1990). Synthesis of research on grade retention. Educational Leadership, 47, 84-88.

TWU Reading Recovery Staff. (1995). Reading Recovery report 1988-95. Denton, TX: Texas Women's University.

Wasik, B.A. & Slavin, R.E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. Reading Research Quarterly, 28, 179-200.

Yin, R.K. (1989). Case Study Research: design and methods. Newbury Park, CA: Sage.